



# Concept of dealing with uncertainty in radar-based data for hydrological purpose

J. Szturc, K. O?ródka, A. Jurczyk, L. Jelonek

## ► To cite this version:

J. Szturc, K. O?ródka, A. Jurczyk, L. Jelonek. Concept of dealing with uncertainty in radar-based data for hydrological purpose. Natural Hazards and Earth System Sciences, 2008, 8 (2), pp.267-279. hal-00299505

**HAL Id: hal-00299505**

**<https://hal.science/hal-00299505>**

Submitted on 27 Mar 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Concept of dealing with uncertainty in radar-based data for hydrological purpose

J. Szturc<sup>1</sup>, K. Ośródka<sup>1</sup>, A. Jurczyk<sup>1</sup>, and L. Jelonek<sup>2</sup>

<sup>1</sup>Institute of Meteorology and Water Management, Radar Centre, Waszawa, Poland

<sup>2</sup>Institute of Meteorology and Water Management, Office of Hydrological Forecasts, Wrocław, Poland

Received: 4 October 2007 – Revised: 14 December 2007 – Accepted: 14 February 2008 – Published: 27 March 2008

**Abstract.** Precipitation radar-based data constitute essential input to Numerical Weather Prediction (NWP) and rainfall-runoff models, however the data introduce a number of errors. Thus their uncertainty should be determined to provide end-users with more reliable information about forecasts. The common idea is to use Quality Index (*QI*) scheme for some number of quality parameters on the assumption that: (1) relationship between the parameters and relevant quality indexes is linear; (2) averaged *QI* is a weighted average of all particular indexes. The uncertainty parameters can be topography-dependent, resulting from spatial and temporal distribution of data, etc. Uncertainty in radar-based data is described by gamma PDF of precipitation, and it is proposed to determine the probability density function (PDF) parameters basing on *QI* values. Practically, precipitation is presented as ensemble of quantiles of the PDF and such an ensemble can constitute input to rainfall-runoff modelling. Since the ensemble is a precipitation input, the hydrological model needs to be activated according to a number of input members.

## 1 Introduction

Radar-derived precipitation estimates and radar-based precipitation nowcasts constitute essential input to Numerical Weather Prediction (NWP) and rainfall-runoff models. However the radar data and nowcasts introduce into final products a number of errors from different sources. Thus their uncertainty should be determined to provide end-users with probabilistic forecasts as more reliable information about rain rate and accumulation. Therefore investigation of data quality and uncertainty in the whole processing chain is necessary.

In literature reviews of weather radar data errors and methods for their estimation are available. Due to the problem of

complexity it is impossible to obtain sufficient information and compute quantitatively the impact of each error. Therefore it seems suitable to evaluate the radar-based data quality not by giving definition and estimation of all radar errors, but by analysing the data properties, among others statistical ones.

The methodology of dealing with radar precipitation uncertainty can be based on a concept of quality index (*QI*) field which is a measure for data quality. Various schemes for creating such a *QI* field for precipitation estimate fields are proposed. An example can be the DLR (Deutsches Zentrum für Luft und Raumfahrt) scheme (Friedrich et al., 2006). The schemes take under consideration main sources of radar limitations significantly affecting the accuracy of radar reflectivity measurements, like beam broadening, ground clutter contamination, attenuation by hydrometeors, and inhomogeneity in vertical profile of reflectivity. This technique requires selecting the most crucial uncertainty parameters to compute quality index fields for radar-based data.

The radar-based precipitation data is the most crucial input for rainfall-runoff modelling, especially for very small mountainous catchments, like the South of Poland, where rainfall-caused flash floods are the most dangerous hydrological events. It is due to required high spatial and temporal resolutions, and short lead-time of forecasts (nowcasts). Therefore the radar-based inputs, their uncertainties, and propagation to hydrological modelling are a subject of this paper.

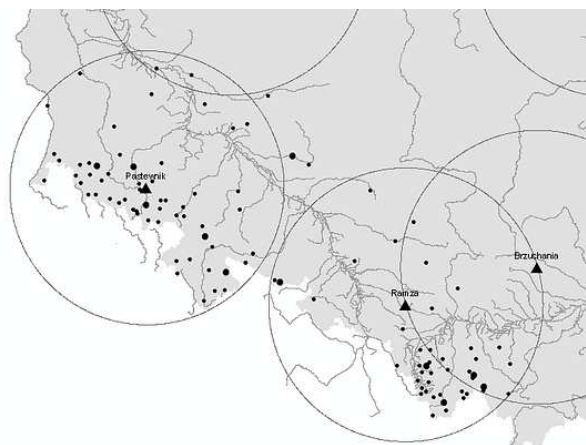
## 2 Test bed

### 2.1 Area

In Poland the Institute of Meteorology and Water Management (IMWM) is responsible for a national meteorological and hydrological service. The IMWM collects data from telemetric (meteorological and hydrological) and remote-sensing networks (weather radars, satellite, etc.).

---

Correspondence to: J. Szturc  
(jan.szturc@imgw.pl)



**Fig. 1.** Map of research area: south of Poland area with rain gauge and weather radar locations.

An area interesting from hydrological point of view is a mountainous region in the south of Poland where the upper Vistula (Wisła) and Odra Rivers are the main sources of flood hazard. In this area a small mountainous catchment was chosen for case study, the Biała Łądecka River catchment of 316 km<sup>2</sup> and altitude in range from 330 up to 1425 m a.s.l.

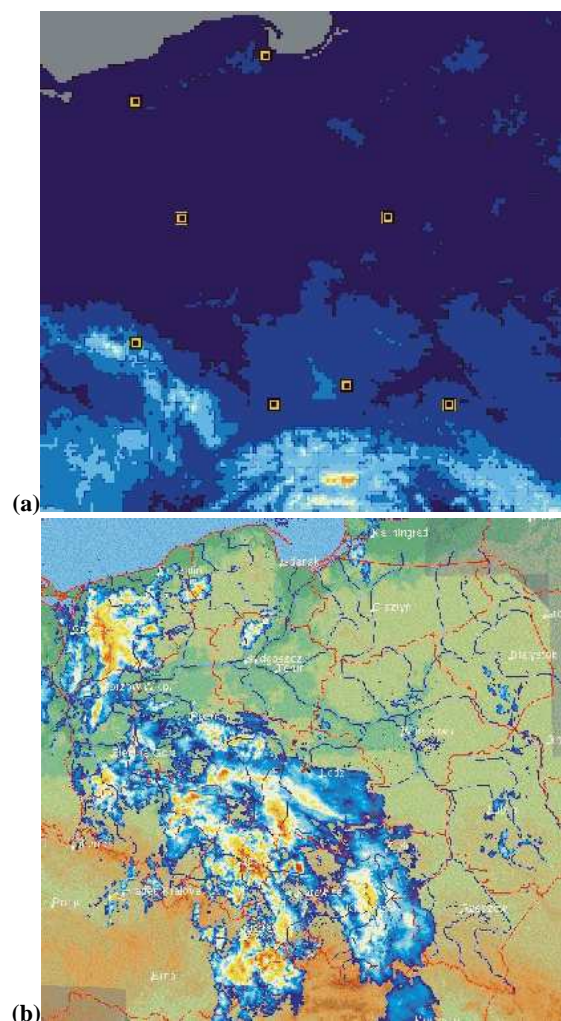
## 2.2 Precipitation data

Telemetric rain gauge data ( $G$ ) as 1-h accumulations from 107 telemetric rain gauges within area shown in Fig. 1 were employed.

Radar data are provided by Polish weather radar network POLRAD that consists of 8 C-Band Doppler radars (Szturc and Dziewit, 2005). They are Gematronik radars with Rainbow software, operated by IMWM (Fig. 2). Among products generated every 10 min the PAC (Precipitation Accumulation) composite is used to deliver 1-h accumulations. The PAC is produced from SRI (Surface Rainfall Intensity) products that measure precipitation on constant height above the ground (in this case 0.7 km).

NIMROD system is UK Met Office software to process radar data using other measurement sources and then to produce analyses and forecasts of precipitation (Weipert and Pierce, 2003). The radar data used as a starting point in NIMROD are provided every 10 min in form of four PPI (Plan Position Indicator) scans at low elevations. The following corrections are applied: ground clutter and anaprop removal, Vertical Profile of Reflectivity (VPR) correction, and Mean Field Bias correction. Corrected radar data are blended with information from other measurement sources, such as ground stations and satellite, to produce NIMROD analyses every 30 min (Golding, 1998).

Hydrological simulations were performed using combination of rain gauge and NIMROD data as precipitation input. The technique of combination is a Sinclair-Pegram method



**Fig. 2.** Polish radar network POLRAD: (a) radar locations, (b) example of radar composite over geographical map from 31 March 2006, 05:30 UTC.

(Sinclair and Pegram, 2005). In this approach the information from radar is used to obtain the correct spatial structure of the precipitation field, while the field values are fitted to the rain gauge observations. The method consists of the following steps: (1) spatial interpolation of rain gauge-derived precipitation data is performed using one of known methods; (2) next errors of the interpolation are determined as differences between two maps: radar data and radar data after spatial interpolation from pixels only with rain gauge locations; (3) finally spatially interpolated rain gauge data is corrected using the differences (errors of interpolation). For spatial interpolation Inverse Distance Weighting method was chosen.

This kind of precipitation field was found as the most proper for this catchment from various kinds of the fields. Quality of the estimation was evaluated (1) by comparison to rain gauge data (using RMSE, bias, correlation coefficient and parameters based on contingency table) and (2) through

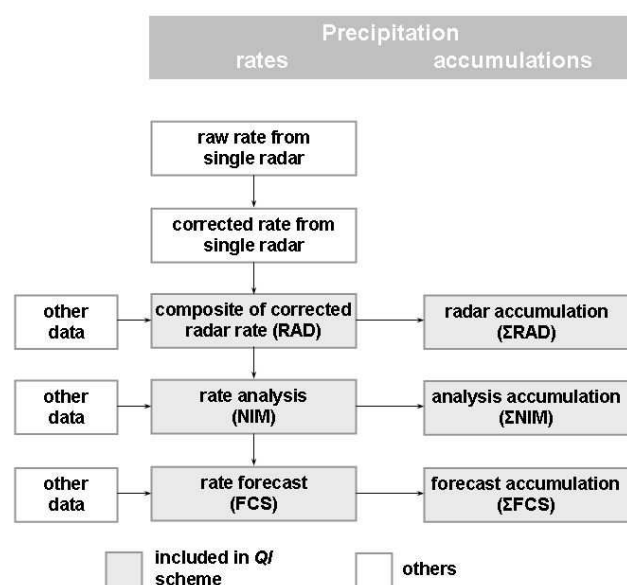


Fig. 3. Radar-based precipitation data.

hydrological analyses of quality of hydrograph simulation from different precipitation inputs (using as criteria among others Nash-Sutcliffe coefficient, correlation coefficient and contingency table parameters) (see report on Task 4.08 of the RISK-AWARE project: Szturc et al., 2007b).

The main idea of the nowcasting philosophy adopted in NIMROD is merging of an extrapolated rainfall analysis with output from a NWP model. NIMROD nowcasts are generated every 30 min with 15-min temporal and 4-km spatial resolution up to six hours ahead.

The proposed Quality Index (*QI*) scheme is to estimate the quality of all kinds of radar-based precipitation data that are generated during the data processing chain. Precipitation data from weather radar POLRAD network are corrected and nowcasted by the NIMROD system using data from other sources.

Generally the processing is carried out in the following steps:

- radar data corrections (ground clutter removal, VPR correction, raingauge adjustment, etc.),
- precipitation field analyses i.e. estimation of radar data complimented with data from other sources (satellite, synoptic stations, last forecasts, etc.),
- precipitation forecasting.

In the presented paper the outline of quality scheme based on a quality index (*QI*) approach is proposed. Six sets of parameters of the scheme can be distinguished for different kinds of precipitation data (Fig. 3):

- composite corrected radar rate (RAD) and its accumulation ( $\Sigma$ RAD),

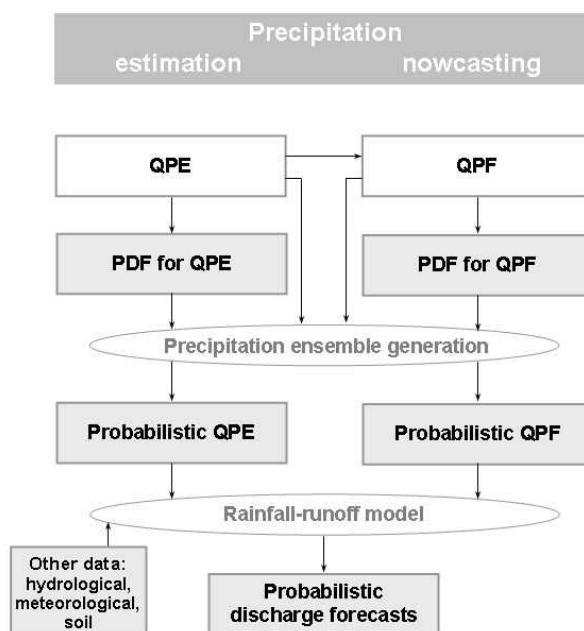


Fig. 4. Scheme of the probabilistic hydrological forecasting.

- rate analysis (from NIMROD or after further processing, e.g. combination with raingauge data) (NIM) and its accumulation ( $\Sigma$ NIM),
- rate forecast (FCS) and its accumulation ( $\Sigma$ FCS).

All employed precipitation data were gathered during rainy events in August 2006.

### 2.3 Hydrological data and rainfall-runoff modelling

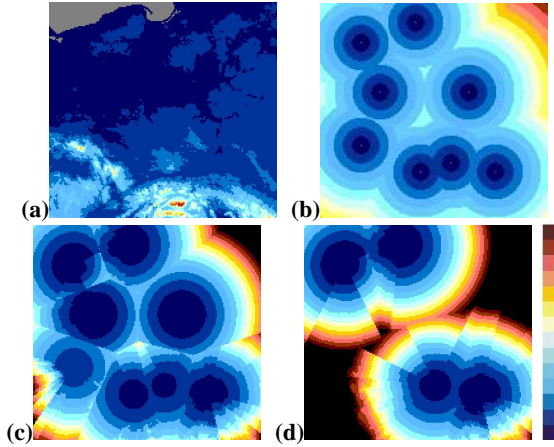
A hydrological test of the proposed techniques for particular precipitation fields was performed on LISFLOOD model (de Roo et al., 2000), that works operationally in Hydrological Forecast Office of IMWM in Wrocław. It is a distributed ( $1 \times 1$  km) physically-based rainfall-runoff model taking account of the influence of topography, precipitation amounts and intensities, antecedent soil moisture content, land use and soil type.

The LISFLOOD model was calibrated on two events: 10–15 August 2006 and 30 August–3 September 2006. A validation was performed on data from 6–9 August 2006 as this period included intense precipitation which triggered the flood event. The research was carried out on selected Biała Łądecka catchment.

## 3 Quality scheme for radar-based precipitation

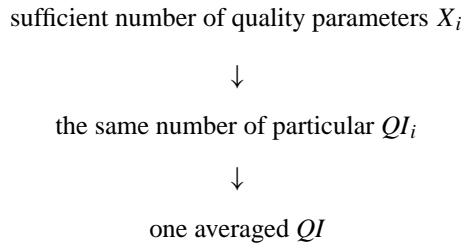
### 3.1 Base

Quality Index (*QI*) schemes become more and more popular in quantitative estimation of precipitation data quality. The



**Fig. 5.** Topography-dependent maps of (a) DEM, (b) DR, and (c, d) MH parameters for Polish radar network. The MH map is shown for two cases: when all radars (c) and only half of them (d) are running. Scale: (a) 0–2.5 km; (b) 0–400 km; (c, d) 0–4 km.

*QI* approach required our knowledge about error sources and their impacts on the data quality. The idea is based on the following chain:



where:  $X_i$  (where  $i=1, \dots, N$ ) are the quality parameters, i.e. specific parameters related to data quality, like technical parameters of radar, quantities connected to distance from radar, spatial and temporal variability of precipitation field, etc.;  $QI_i$  are the quality indexes for relevant  $X_i$  parameters;  $QI$  is the averaged quality index.

The common idea is to use the scheme for a few quality parameters assuming that (Friedrich et al., 2006; Szturc et al., 2006a; 2007a):

1. The relationship between  $X_i$  and  $QI_i$  is linear:

$$QI_i = \begin{cases} 1 & \text{for } X_i \leq X_{i0} \\ 0 & \text{for } X_i \geq X_{i1} \\ \frac{X_{i0}-X_i}{X_{i0}-X_{i1}} & \text{for } X_{i0} < X_i < X_{i1} \end{cases} \quad (1)$$

where  $X_{i0}$  and  $X_{i1}$  are the boundary values for the  $X_i$  parameter. For some quality parameters the higher parameter, the higher quality index is (e.g. number of products that are incorporated in a given accumulation). In this case in the relationship (1) the inequality signs must be replaced by opposite ones.

2. The averaged  $QI$  is calculated as a weighted average of all  $QI_i$  indexes:

$$QI = \sum_{i=1}^N QI_i \cdot W_i \quad (2)$$

where  $W_i$  are the weights of particular  $QI_i$  indexes. The weights can be estimated basing on investigation of historical data as it is described in Sect. 5.

The additional assumption that if any  $QI_i$  is over a critical value  $X_{icrit}$  then averaged  $QI$  should be taken as equal 0 is recommended here. It is to save the scheme from producing quite wrong data.

The above assumption about linearity of  $X_i$  and  $QI_i$  relationship is made as it is difficult to find exact form. However some parameters  $X_i$  from described in Sect. 4 does not show such a behaviour, for instance increase of distance to radar site results in nonlinear quality decrease.

### 3.2 Concept of precipitation uncertainty propagation into hydrological modelling

The presented methodology of dealing with radar precipitation uncertainty propagation to runoff forecasts is based on a concept of (Fig. 4):

1. quality index field for all kinds of precipitation data,
2. PDF that is employed to characterise the phenomenon,
3. parameterisation of the PDF.
4. ensemble of the PDF quantiles as input to rainfall-runoff models,
5. ensemble of runoff forecasts.

## 4 Quality parameters

### 4.1 Main radar errors

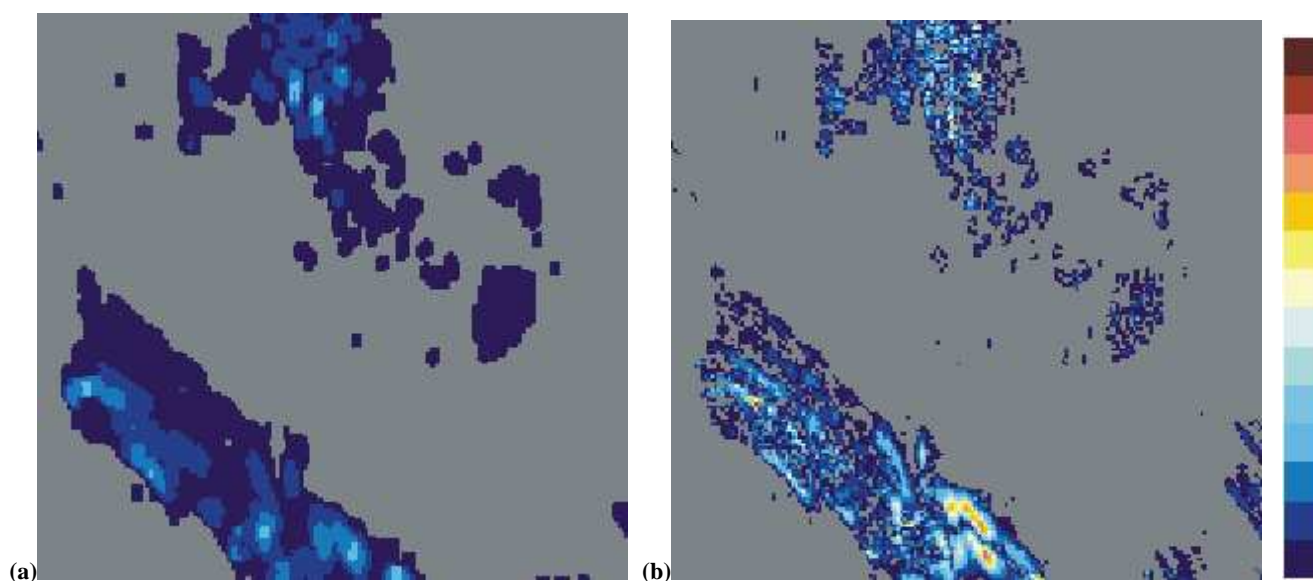
The selected quality parameters, which are to characterise all kinds of data, have been divided into particular groups.

Two major kinds of uncertainty parameters are employed: (1) topography-dependent and (2) resulting from spatial and temporal distribution of radar-based data. For real-time application it should be remembered that the particular quality parameters must be easily determinable.

For quality of precipitation accumulation a number of rain-rate products that compose a given sum is important and is introduced as an additional quality parameter.

Two next parameters incorporated into the proposed scheme are used only for forecasts. They are the lead-time (*LT*) of forecasts and quality index previously determined for estimated precipitation (*QIE*).





**Fig. 6.** Example of variability-dependent maps of radar data: **(a)** SV and **(b)** TV parameters for Polish radar network. Hourly accumulation generated on 14 August 2006, 03:00 UTC. Scale: 0–10 mm.

**Table 1.** Main radar errors and related quality parameters.

Error	Quality parameter	Magnitude	Frequency	Range
Hardware problems, miscalibration, pointing error, etc.	–	medium – big	continuous	continuous
Earth curvature	<i>DR, MH</i>	medium	continuous	continuous
VPR variability	<i>DR, MH</i>	medium	seasonal	continuous
Spatial resolution	<i>DR, MH</i>	small	continuous	continuous
Beam blocking, shielding	<i>MH</i>	big	continuous	local
Total beam overshooting	<i>MH</i>	small – big	seasonal	local
Ground clutter	<i>DEM</i>	small – big	continuous	local
AP clutter, propagation changes	–	small – medium	seasonal	local
Interfering emitters, jamming	–	small – medium	occasional	local
Attenuation by precipitation	<i>DR</i>	small	continuous	continuous
Attenuation by wet/icy radome	–	small	seasonal	local
Hail, water phase, $Z - R$ relationship	<i>SV</i>	small	seasonal	continuous
Orographic enhancement	<i>DEM</i>	small – medium	continuous	local
Overhanging precipitation	–	medium	seasonal	local
Temporal resolution	<i>NP, TV</i>	medium	continuous	continuous

Any number of other parameters can be added.

In Table 1 the most important error factors and parameters that can describe the risk level of burdening with these errors are listed. The list is valid especially for POLRAD network associated with NIMROD system. It is not feasible to produce the list which would be strict and unambiguous. The list was prepared basing on experiences of Radar Centre staff in Institute of Meteorology and Water Management and other similar attempts, especially: Šálek et al., 2004; Michel-

son et al., 2005; NORDRAD (*Future...*, 2003). In Table 2 a similar list for forecast quality factors is shown.

## 4.2 Definitions of quality parameters

### *Topography-dependent parameters*

Distance from radar (*DR*) is a crucial parameter as radar beam expands with distance from the radar site. Moreover the curvature of the Earth surface results in increasing the

**Table 2.** Main forecasting errors and related quality parameters.

Error	Quality parameter	Magnitude	Frequency	Range
Quality of starting data	<i>QIE</i>	big	continuous	continuous
Lead-time	<i>LT</i>	big	continuous	continuous
Spatial resolution	<i>SV</i>	medium	continuous	continuous
Temporal resolution	<i>NP, TV</i>	medium	continuous	continuous

vertical distance of the radar beam to the ground. For the lowest scan the beam height above the ground can reach over 2.5 km at a 200-km distance from radar (Fig. 5c, 5d).

The next parameter is a spatial pattern of altitude (*DEM*) that is represented by DEM (digital elevation map). The important effects resulting from the presence of higher mountain peaks are radar data contamination by ground clutters, radar beam blocking, and shielding (Fig. 5a).

Third topography-dependent parameter is a height of the lowest scan (*MH*). This height indicates the lowest point of vertical profile of atmosphere that can be visible by any radar beam since the lower altitudes are shielded by terrain between radar site and a given location. It can be seen that this parameter combines both previously mentioned (Fig. 5a, 5b, 5c). The *MH* field differs from *DR* mainly in mountainous area (in Poland it is in the South of the country, see Fig. 2a).

These three parameters are dependent on each other, so two of them can be chosen as uncertainty parameters in the proposed quality index scheme.

#### *Spatial and temporal variability of precipitation fields*

Next, the spatial (*SV*) and temporal (*TV*) variability of data fields are investigated and used as measures of data uncertainty.

Spatial variability (*SV*) is determined for each pixel on grid bigger than radar pixel size. The grid can be e.g. 3×3 or 5×5 pixels (in our case), i.e. of size 12 or 20 km, if the pixel size is 4 km. The measure of the variability is variance (Fig. 6a).

Temporal variability (*TV*) is determined for each pixel on *n*-hour moving time-window (where *n*=6 in our scheme). If a 1-h accumulation is input then the six previous time steps are taken in case of QPE and six subsequent forecasts in case of QPF. The measure is variance as well (Fig. 6b).

#### *Representativeness of accumulations*

An uncertainty of precipitation accumulation estimate is strongly dependent on a number of rain rate products (*NP*) incorporated in particular hourly accumulation (from 0 to 7 maps maximally). The number results from the actual availability of the data during the given hour and the POLRAD network and NIMROD time-steps of data production.

#### *Lead-time of forecasts*

The uncertainty depending on lead-time (*LT*) is determined directly from the lead-time value expressed in hours. For forecasts *LT* varies from 1 to 6 h as maximally 6-h radar-based nowcasts are generated in NIMROD system. The same values of *LT* are taken for the whole map.

#### *Quality index for estimated precipitation*

The quality of advection forecasts (nowcasts) strongly depends on *QIE* – the quality of the initial field i.e. an estimate of precipitation. Therefore, the averaged quality index *QI*, calculated from formula (2) for that estimate, constitutes an important factor *QI<sub>QIE</sub>* in forecast quality as well. Then the field of averaged *QI* for estimate is taken as the last parameter of forecast quality.

#### 4.3 Quality indexes *QI<sub>i</sub>* and averaged quality index *QI*

The *QI<sub>i</sub>* fields calculated on the base of the parameters *X<sub>i</sub>* using Eq. (1) are fixed for a given radar location and are independent of weather situations in the case of *LT*, *DR*, and *MH*. Since composite maps are used instead of single site radar ones, the *DR* and *MH* maps (Fig. 5c, 5d) cannot be fixed when not all radars operate at given time. The *QI<sub>i</sub>* fields based on *SV* and *TV* (Fig. 6) are variable temporally and spatially so they have to be calculated for each time step separately.

Finally Eq. (2) can be written in the following form:

$$\begin{aligned}
 QI &= \sum_{i=1}^N QI_i \cdot W_i \\
 &= W_{DR} QI_{DR} + \\
 &\quad + W_{MH} QI_{MH} + W_{SV} QI_{SV} + \\
 &\quad + W_{TV} QI_{TV} + W_{LT} QI_{LT} + \\
 &\quad + W_{NP} QI_{NP} + W_{QIE} QI_{QIE}
 \end{aligned} \quad (3)$$

with condition for the weights:

$$W_{DR} + W_{MH} + W_{SV} + W_{TV} + W_{LT} + W_{NP} + W_{QIE} = 1 \quad (4)$$

The Eq. (3) is valid for both estimate and forecast, with some weights equal to zero in dependence on the kind of data.

**Table 3.** Correlations between some quality parameters and  $\ln(R/G)$  for 1-h and 24-h accumulations. Radar PACs are used as  $R$  data; precipitation threshold is 0.5 mm.

Quality parameter	Correlation	
	1 h	24 h
<i>DR</i>	0.325	0.357
<i>DEM</i>	−0.053	−0.171
<i>MH</i>	0.306	0.375
<i>SV</i>	−0.007	0.635
<i>TV</i>	0.127	0.608

## 5 Determination of the quality parameter weights

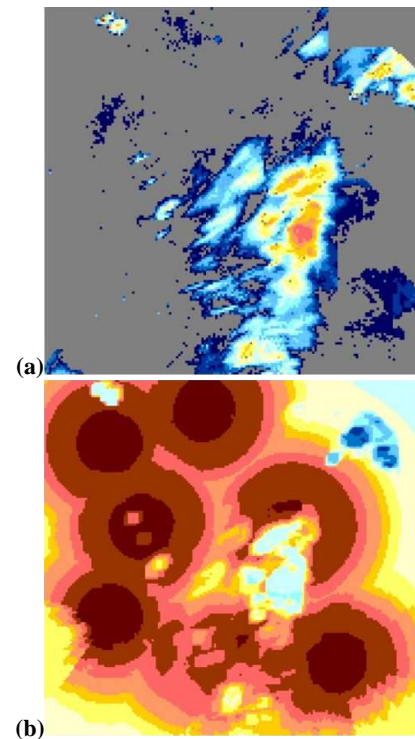
In the proposed scheme quality index  $QI_i$  for each quality parameter  $X_i$  is calculated assuming a linear relationship between them. This calculation is made for each radar-map pixel and in this way the field of the quality index is obtained. Next all the individual fields are summarised to an averaged  $QI$  field using appropriate weights (Eqs. 3 and 4). The scheme should be calibrated in order to get objective information about data quality. It requires a procedure for the parameterisation of the scheme that involves determining some quantities for each parameter  $X_i$ : its weight and lower  $X_{i0}$  and upper  $X_{i1}$  thresholds. The values are defined as follows:  $X_{i1}$  is boundary value  $X_i$  for which quality index  $QI_i = 1$ , and  $X_{i0}$  is boundary value  $X_i$  for which  $QI_i = 0$ , so beyond these thresholds the quality index  $QI_i$  values are set to 0 or 1. Moreover a critical value ( $X_{icrit}$ ) is introduced, which means that if the parameter  $X_i$  reaches the  $X_{icrit}$  value then the averaged  $QI_i$  value is set to zero for the pixel even if other parameters are quite good.

The parameterisation of the proposed scheme is performed on a historical dataset. It is assumed that raingauge data are exact in their locations therefore information about differences or the ratio between raingauge ( $G$ ) and radar ( $R$ ) observations in these points can be a measure of quality. The correlation between given quality parameter  $X$  and measured errors, represented by e.g.  $\ln(R/G)$  or  $(R - G)$ , can indicate the parameter importance in terms of data quality (here  $R$  means all radar-based data).

Correlations computed for August 2006 with employed quality parameters are listed in Table 3.

For comparison correlations between  $R$  and  $G$  are 0.311 for 1-h accumulations and 0.523 for 24-h accumulations.

Basing on these correlations the optimal weights in Eq. (2) can be determined for particular  $QI_i$  fields (Table 4). Moreover threshold and critical values for linear interpolation of these parameters are determined and finally averaged  $QI$  is computed from Eq. (2). The quality information field obtained in this way is attached to the radar-based precipitation product (see Fig. 7).



**Fig. 7.** Example of: (a) radar data and (b) assigned averaged  $QI$  field for  $\Sigma$ RAD data: 1-hour accumulation from 15 August 2006, 08:00 UTC, POLRAD network. Scale: (a) 0–10 mm; (b) 0–1.

An example of  $QI$  map is shown in Fig. 7. It can be noticed that there is a good data quality where there is no rain, and a reduced quality for wet pixels. It can be explained in the following way. Estimation  $R=0$  is burdened with uncertainty, but usually not very significantly, as only low-intensity rain may be expected to occur in these places. However if the rain is measured then the uncertainty is higher because of variability of rain field. In general: the higher the rain, probably the higher the uncertainty is as well.

The most crucial task in this  $QI$  scheme is setting the thresholds  $X_{i0}$ ,  $X_{i1}$ , and  $X_{icrit}$ . They should be determined basing on physical features of particular parameters. For example a distance range to get the proper data from weather radar can be estimated approximately. It is advised to take into account e.g. data application, a distinction between stratiform and convective precipitation, etc.

## 6 Precipitation PDF

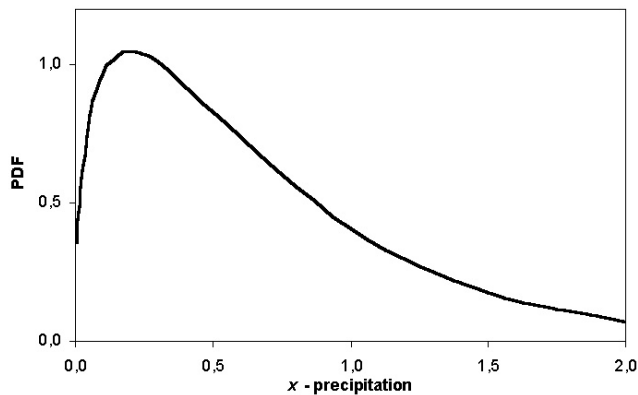
### 6.1 Precipitation ensemble – a review

The technique of ensemble of precipitation fields is commonly employed considering uncertainty in the NWP data. However NWP-derived precipitation forecasts can only be used in applications where high resolution is necessary. For



**Table 4.** Example of quality scheme parameters for radar-based precipitation data.

Quality parameter $X_i$	Unit	$X_{i1} (QI_i=1)$	$X_{i0} (QI_i=0)$	$X_{icrit}$	Weights for ...					
					RAD	$\Sigma$ RAD	NIM	$\Sigma$ NIM	FCS	$\Sigma$ FCS
Distance to nearest radar ( <i>DR</i> )	km	10	100	>200	0.180	0.180	0.180	0.180	0	0
Min. Height ( <i>MH</i> )	km	0.5	5	>5	0.190	0.190	0.190	0.190	0	0
Spatial variability ( <i>SV</i> )	mm	0.01	0.1	>100	0.322	0.322	0.322	0.322	0.322	0
Temporal variability ( <i>TV</i> )	mm	0.001	0.1	>100	0.308	0	0.308	0	0.308	0
Number of products ( <i>NP</i> )	–	7	3	$\leq 2$	0	0.308	0	0.308	0	0.500
<i>QIE</i>	–	1	0	<0.1	0	0	0	0	0.185	0
Lead-time ( <i>LT</i> )	hour	0	7	$\geq 7$	0	0	0	0	0.185	0.500

**Fig. 8.** Example of gamma PDF ( $p=2.4$ ,  $b=2.0$ ).

instance COSMO-LM 14-km spatial resolution is too poor for small mountainous catchments (of 100–1000 km<sup>2</sup> area) where flash floods are important events. A common approach for dealing with uncertainty in NWP-derived forecasts is to determine an exceedance probability of forecasted precipitation.

This can be achieved by working on ensembles of forecasts, that are implemented as ensemble prediction system (EPS), e.g. in ECMWF EPS (Roulin and Vannitsem, 2005). It consists of 50 perturbed members (to 10 days with 80-km resolution every 24 h) and one unperturbed forecast (with 40-km resolution) (Molteni et al., 1996). These EPS forecasts can be employed only for bigger catchments. The perturbation technique may be Monte Carlo sampling (e.g. Pappenberger et al., 2005), also different NWP forecasts can be used, e.g. including different resolutions from the same model.

Such an ECMWF EPS ensemble can be used either to calculate quantiles that constitute input to rainfall-runoff modelling, or all the members are directly used if the calculation time is not critical as 51 deterministic flood forecasts are generated (Gouweleeuw et al., 2005).

A different approach is based on the determination of a probabilistic QPF (so called PQPF) like it is done in Na-

tional Weather Service Forecast Office in Tulsa, Oklahoma (Amburn, Frederick, 2006). That is derived from two quantities: deterministic QPF, that is treated as mean precipitation amount, and Probability of Precipitation (PoP). The PoP is calculated in dependency on present ground weather reports. It is assumed that PQPF in means of Probability of Exceedance (PoE) can be approximated by the exponential probability density function (PDF). The description above is for NWP fields. In the case of radar precipitation, the production of the ensembles is more difficult but such a concept becomes more and more popular nowadays. It is proposed to introduce disturbances into the radar data (Germann et al., 2006).

The availability of an ensemble of precipitation estimates and forecasts opens up the possibility of constructing multiple runoff forecasts and using them to quantify the accuracy of the forecasts and the likelihood of a warning threshold being exceeded. The uncertainty in the precipitation input is likely to dominate runoff forecasts accuracy, then other sources of errors, such as rainfall-runoff model itself errors can be ignored here. The precipitation uncertainty information can constitute a first guess in investigation of hydrological forecast uncertainty.

## 6.2 Gamma PDF

It is proposed to take into account the uncertainty in estimates or forecasts of precipitation using a specific PDF suitable to reflect physical features of rainfall. The gamma distribution can be used for this purpose (e.g. Amburn and Frederick, 2006, 2007):

$$PDF(x) = \frac{b^p}{\Gamma(p)} x^{p-1} e^{-bx} \quad (5)$$

where  $p$  and  $b$  are the PDF parameters,  $p, b > 0$ ;  $\Gamma$  is the gamma function (see Fig. 8).

Radar measurements are treated as random variable  $R$ . A specific value of  $R$  will be denoted as  $r$ . The true value of precipitation is searched, that is approximately taken as a raingauge value  $G$  (if it was available) in a pixel containing it. It is needed to know the true values in each pixel and their uncertainties.

A new random variable  $X$  that represents the true precipitation is introduced:

$$X = E(G|R = r) \quad (6)$$

that means expectation  $E$  of the true value  $G$  under condition that radar value  $R$  equals  $r$ .

In the case of a continuous variable  $R$  the interval of width  $\varepsilon$  (increasing with  $r$ ) is taken instead of one value  $r$ :

$$X = E(G|R \in [r - \varepsilon, r + \varepsilon]) \quad (7)$$

Statistical characteristics of the  $X$  variable can be estimated from other, more available data. It is assumed that the variance:

$$\text{var}(X) = \text{var}(R - G) \quad (8)$$

Another measure of the  $R$  field uncertainty can be applied, like  $\log(G/R)$ , instead of  $(R - G)$ .

If the last relationship is true, then the variance  $\text{var}(R - G)$  can be estimated on historical data using radar  $R$  and raingauge  $G$  data together with appropriate  $QI$  data.

To estimate the PDF for radar data it is assumed that:

1. the PDF parameters are related to the averaged  $QI$  value,
2. this relationship can be experimentally determined for each pixel of the precipitation data field.

As a consequence the probabilistic precipitation field may consist of three values for each pixel (see Fig. 9): two PDF parameters (or more in dependence on the specific PDF) and QPE (or QPF).

### 6.3 Relationship between $\text{var}(R - G)$ and $QI$ determined on historical dataset

The  $QI$  fields are computed for radar-based precipitation fields using Eq. (3) and appropriate weights. Having these fields the next step is to determine a precipitation PDF (which can be gamma PDF from Eq. 5) for each pixel of the fields. In order to solve this task the parameterisation of the PDF is required, i.e. calculation of both PDF parameters.

The concept proposed in the paper is to determine the variance  $\text{var}(R - G)$  that is the variance of the radar-raingauge difference observed at raingauge locations on historical data. Then the linear regression of the relationship between this variance  $\text{var}(R - G)$  and the quality index  $QI$  is to be computed. In practice it may be done as follows.

If the following historical data are available: radar/NIMROD data  $R$ , raingauge data  $G$  (both as e.g. hourly accumulations), and  $QI$  field, then at first  $QI$  values (that vary in range from 0 to 1) are divided into a number  $n$  of classes (for instance  $n=10$ ). Then the variance of  $(R - G)$  is calculated for each class:

$$\text{var}(R - G) = \frac{1}{N} \sum_{i=1}^N \left( (R_i - G_i)^2 - B^2(R, G) \right) \quad (9)$$

where  $N$  is the number of radar-raingauge pairs;  $B(R, G)$  is defined as follows:

$$B(R, G) = \frac{1}{N} \sum_{i=1}^N (R_i - G_i) \quad (10)$$

Then relationship between  $\text{var}(R - G)$  and  $QI$  is estimated using linear (or any non-linear) regression. As a result an interpolated variance  $\overline{\text{var}}(R - G)$  is determined:

$$\overline{\text{var}}(R - G) = a_1 \cdot QI + a_2 \quad (11)$$

where  $a_1$  and  $a_2$  are the linear regression coefficients.

In Fig. 10 an example of the relationship between  $\text{var}(R - G)$  and  $QI$  divided into classes is shown. For this example the relationship (Eq. 11) was established as:  $\overline{\text{var}}(R - G) = -1.18 \cdot QI + 0.98$ , with a correlation coefficient equal to 0.92. The data were NIMROD analyses (i.e. processed radar data  $\Sigma NIM$ ) that seem to be the best NIMROD precipitation estimation (Szturc et al., 2007b). The data were combined with raingauge data for selected flood event from 6–9 August 2006. For comparison: using radar-derived PAC (Precipitation Accumulation) data the relationship was established as:  $\overline{\text{var}}(R - G) = -4.46 \cdot QI + 4.81$ , with a correlation coefficient equal to 0.85.

### 6.4 Real-time relationship between the PDF parameters and both precipitation $R$ and quality index $QI$

Having this last relationship between the variance  $\text{var}(R - G)$  and quality index  $QI$  computed on historical data, it is possible to estimate in real-time the following statistical moments of precipitation for a particular precipitation field: expectation  $E(X)$  and variance  $\text{var}(X)$ . Now the parameters  $p$  and  $b$  are computed from  $E(X)$  and  $\text{var}(X)$  values as it is described below.

Gamma  $PDF(x)$  is defined by Eq. (5). Relationships between the PDF parameters  $p$  and  $b$ , and precipitation statistical moments  $E(X)$  and  $\text{var}(X)$  are as follows:

$$E(X) = \frac{p}{b}; \text{var}(X) = \frac{p}{b^2} \quad (12)$$

From the above equation system both parameters  $p$  and  $b$  can be determined:

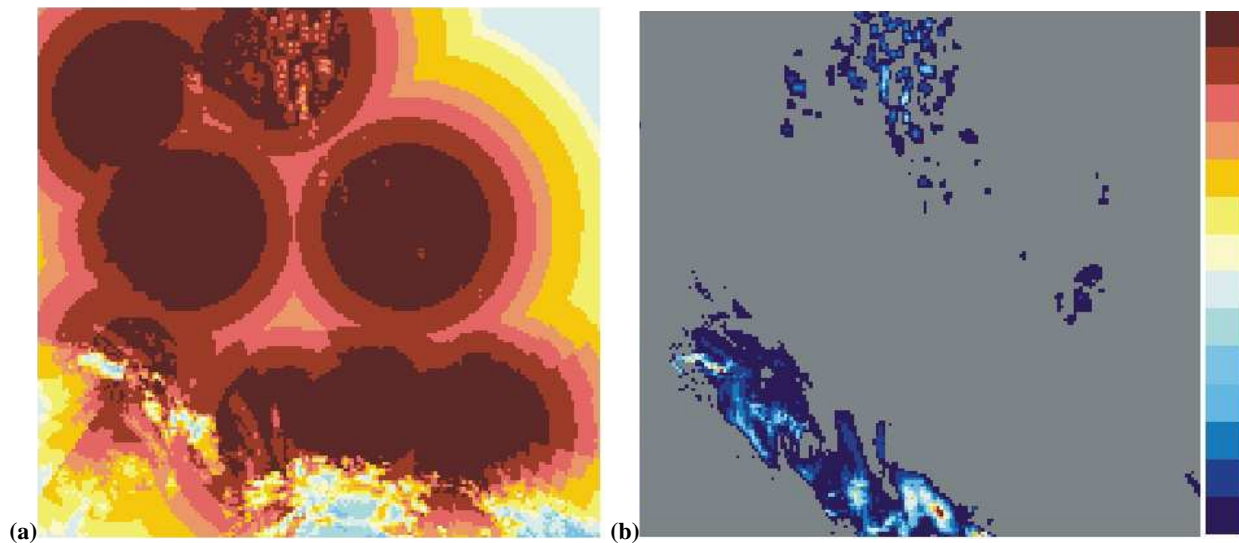
$$p = \frac{E^2(X)}{\text{var}(X)}; \quad b = \frac{E(X)}{\text{var}(X)} \quad (13)$$

In Eq. (13) the expectation  $E(X)$  may be identified with the radar/NIMROD precipitation estimate  $R$ , whereas the variance  $\text{var}(X)$  is calculated from the interpolated relationship between  $\overline{\text{var}}(R - G)$  and  $QI$  (Eq. 11). Therefore:

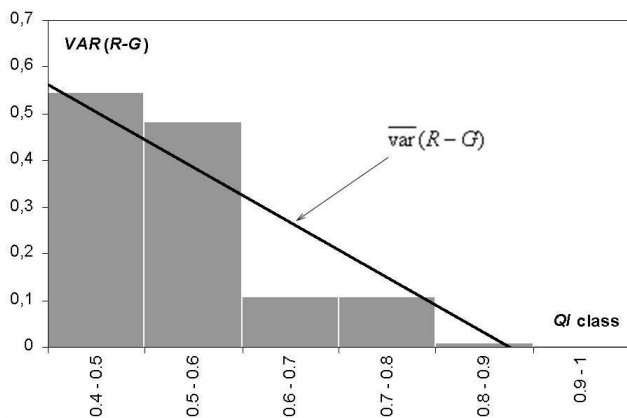
$$E(X) = r; \quad (14)$$

$$\text{var}(X) = \overline{\text{var}}(R - G) = a_1 \cdot QI + a_2 \quad (15)$$

Finally, for a given data pixel with the known  $QI$  and  $R$  values both  $p$  and  $b$  parameters can be calculated from the above equation system (Eq. 14–15). In this way the precipitation PDF can be computed for each data pixel.



**Fig. 9.** Example of maps of gamma PDF parameters: (a)  $p$  and (b)  $b$  for Polish radar network for hourly accumulation generated on 14 August 2006, 03:00 UTC. Scale: 0–10.



**Fig. 10.** Example of determination of relationship between  $\text{var}(R-G)$  and  $QI$  ( $\Sigma$ RAD 1-h data from 6–9 August 2006).

## 6.5 Probabilistic precipitation – ensemble of precipitation quantiles

The rainfall-runoff model is a deterministic model that requires not probabilistic precipitation as its input but deterministic one. In order to overcome this difficulty a common solution is to produce an ensemble of a few deterministic inputs instead of only one. It may be done by selection of some characteristic maps. From a practical aspect a number of members of ensembles should not be too big. The ensemble members can be chosen as quantiles, e.g.  $q\%=5, 25, 50, 75$ , and  $95\%$  basing on a cumulative distribution function CDF. The  $q\%$ -quantiles of the PDF, i.e.  $P_{q\%}$  is calculated

from the gamma CDF:

$$P_{q\%} = F(q\%) = \int_{-\infty}^{q\%} f(x)dx = \int_0^{q\%} \frac{b^p}{\Gamma(p)} x^{p-1} e^{-bx} dx \quad (16)$$

This ensemble is used to compute an ensemble (sequence) of inputs to a deterministic rainfall-runoff model. The way is as follows. First distances between  $P_{50\%}$  and all other quantiles are calculated, i.e.  $P_{5\%}-P_{50\%}$ ,  $P_{25\%}-P_{50\%}$ ,  $P_{75\%}-P_{50\%}$ , and  $P_{95\%}-P_{50\%}$ . Next is to calculate the corrected quantiles according to:

$$R_{q\%} = R + (P_{q\%} - P_{50\%}) \quad (17)$$

where  $R$  is the radar-based precipitation. E.g.  $R_{75\%} = R + (P_{75\%} - P_{50\%})$ , etc.

## 7 Probabilistic runoff

### 7.1 Precipitation ensemble as input to rainfall-runoff model

The proposed hydrological ensemble prediction system is designed not only to forecast the most likely hydrological scenario, but also to estimate the uncertainty of such a forecasting due to the meteorological uncertainty. The uncertainty introduced by rainfall-runoff model is not taken into account here.

Since input to rainfall-runoff model is probabilistic, an ensemble of precipitation fields is provided. In consequence the hydrological model needs to be activated 5 or 7 times according to a number of quantiles (Szturc et al., 2006b). As an output from the rainfall-runoff model the same number of 5



**Fig. 11.** Example of maps (from left to right): P5%, P25%, P50%, P75%, and P95% quantiles of gamma PDF for Polish radar network for hourly accumulation generated on 14 August 2006, 03:00 UTC. Scale: 0–10 mm.

or 7 discharge hydrographs will be obtained (Fig. 12) which define classes of runoff uncertainty.

Classes of runoff probability will be determined basing on the statistical investigation of a big number of rainfall flood events (Gouweleeuw et al., 2005; Roulin and Vannitsem, 2005).

A more advanced way of direct using probabilistic input to deterministic rainfall-runoff model and deriving probabilistic runoff forecast is an analytic-numerical Bayesian forecasting system (BFS) proposed by Krzysztofowicz (Kelly and Krzysztofowicz, 2000; Krzysztofowicz, 2002; Herr et al., 2002). Within a framework of BFS Krzysztofowicz (2002) proposed a methodology to produce a probabilistic river stage forecast for the short range. The system decomposes the total forecasting uncertainty into input uncertainty and hydrological uncertainty. The former source of uncertainty is dominated by the unknown future precipitation; the latter generally aggregates all other uncertainties and includes model, parameter, estimation and measurement errors.

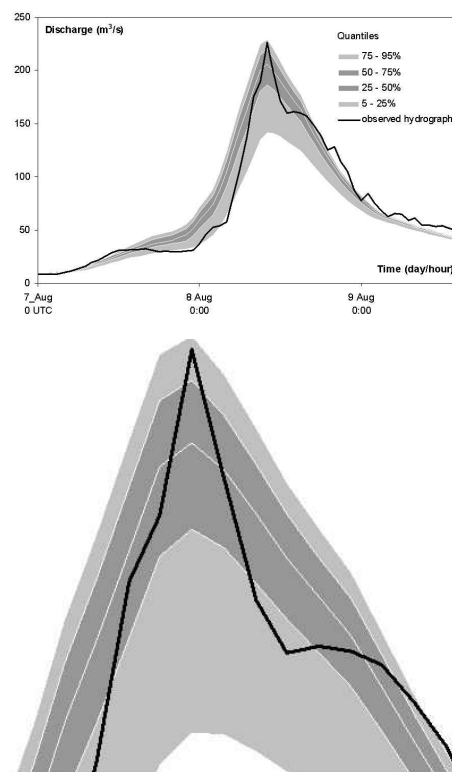
Having the sequence of ensemble  $R_{q\%}$  of precipitation input to rainfall-runoff model, the ensemble of hydrographs (i.e. sequence of discharge  $H_{q\%}$ ) can be obtained as the whole ensemble of precipitation estimates is employed to generate forecasts. That means the rainfall-runoff model should run for each quantile separately.

Next task is to estimate the PDF of discharge for a given time step like it was done for precipitation. However because of the non-linear relationship between precipitation and discharge a simple transformation of uncertainty is not possible.

## 7.2 Example

The example of the hydrograph ensemble is shown in Fig. 12. Data from a 6–9 August 2006 flood event were used here. The used LISFLOOD rainfall-runoff model was calibrated for various kinds of deterministic precipitation inputs and verified on independent data in order to choose the best input. As a result, raingauge-radar combination by Sinclair-Pegram method was found the best one (see Sect. 2.2). The examined event had not been used in the calibration.

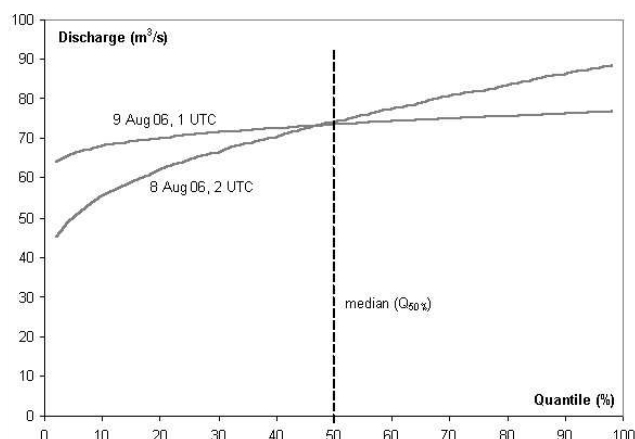
Uncertainty in river runoff forecasting is a result of quality of the whole data set and all particular uncertainties of pro-



**Fig. 12.** Ensemble of probabilistic precipitation fields (5, 25, 50, 75, and 95%) as input to rainfall-runoff model and resulting ensemble of hydrographs and observed hydrograph. On the right an enlarged excerpt of the hydrographs peaks is shown.

cessing chain. The uncertainties in output data are mainly caused by the errors in:

- precipitation measurement (weather radar, raingauge etc.) and estimation (VPR correction etc.),
- precipitation forecasting (radar-based, NWP),
- other meteorological measurement (temperature etc.),
- soil measurement and data (soil type, land use, etc.),
- hydrological measurement (water level, discharge, etc.),
- rainfall-runoff model algorithms.



**Fig. 13.** Diagrams of discharge and quantiles relationship for two specific time steps (8 August 2006, 02:00 UTC and 9 August 2006, 01:00 UTC) with similar median value.

From the above list the first two errors are the most significant and they are introduced into the proposed *QI* scheme.

In Fig. 12 the precipitation uncertainty is visible by the range of the hydrogram ensemble. Since the observed hydrogram is only partially within the ensemble range, it is most probable that it is not because of precipitation errors but rather influence of the *QI* scheme parameterization.

In Fig. 13 two diagrams of relationship between discharges and quantiles for two specific time steps with similar median value are presented. The flatter curve is assigned to higher quality radar data, i.e. *QI* averaged on investigated catchment is more close to one.

The relationships in the range of quantiles above 25% became linear, so transformation between precipitation field quantiles and runoff quantiles may be linear as well. This observation is corresponding to the Horton overland flow: lower rates of rainfall influence the overflow less because of higher infiltration.

## 8 Remarks

The task of the paper was to propose a more common quality index (*QI*) scheme, e.g. for international exchange of radar-based data, that must not be assigned to algorithms used for making specific products for domestic applications.

The proposed scheme is not complicated, but universal since it can be employed for each radar-based data originating from any network and does not matter what measurement techniques and algorithms for data processing have been used. As a result, the implementation of the scheme is easy. However parameterisation of the scheme is ambiguous so it is not an easy task because clear criteria of quality are not known and commonly understood.

In next part of the paper it was showed how to deal with the quality information while using the radar-based data in hy-

drological rainfall-runoff modelling, especially how the input data uncertainty is transformed into hydrological data (river discharge). It is a way of precipitation data uncertainty incorporation into hydrological modelling.

Up to now the quality index scheme has been operated only in a research mode. A precise evaluation of different weighting factors for different applications and weather situations should be done in operational mode.

**Acknowledgements.** This work has been carried out within the framework of the COST-731 Action “Propagation of uncertainty in advanced meteo-hydrological forecast systems” and supported by the Polish Ministry of Science and Higher Education.

Edited by: P. P. Alberoni

Reviewed by: two anonymous referees

## References

- Amburn, S. and Frederick, J.: An experiment in probabilistic quantitative precipitation forecasting. 86th AMS Annual Meeting. Proceedings, 28 January–7 February 2006, Atlanta, GA, (<http://ams.confex.com/ams/pdfpapers/100354.pdf>), 2006.
- Amburn, S. and Frederick, J.: Probabilistic quantitative precipitation forecasting. 22nd Conference on Weather Analysis and Forecasting, 25–29 June 2007, Park City, Utah (<http://ams.confex.com/ams/pdfpapers/124735.pdf>), 2007.
- de Roo, A. P. J., Wesseling, C. G., van Deursen, W. P. A.: Physically based river basin modelling within a GIS: the LISFLOOD model, *Hydrol. Process.*, 14, 1981–1992, 2000.
- Friedrich, K., Hagen, M., and Einfalt, T.: A quality control concept for radar reflectivity and Doppler velocity, *J. Atmos. Oceanic Technol.*, 23, 865–887, 2006.
- Future methodologies for Nordrad and Baltrad, (NORDRAD web page: <http://nordrad.fmi.fi/methods/Problem.class.html>), 2003.
- Germann, U., Berenguer, M., Sempere-Torres, D., and Salvadè, G.: Ensemble radar precipitation estimation – a new topic on the radar horizon, *Proceedings of ERAD 2006*, 559–562, 2006.
- Golding, B. W.: NIMROD: A system for generating automated very short range forecasts, *Meteor. Appl.*, 6, 1–16, 1998.
- Gouweleew, B. T., Thielen, J., Franchello, G., de Roo, A. P. J., and Buizza, R.: Flood forecasting using medium-range probabilistic weather prediction, *Hydrol. Earth Syst. Sci.*, 9, 365–380, 2005, <http://www.hydrol-earth-syst-sci.net/9/365/2005/>.
- Gutiérrez, J. M. and Aguado, F.: Quality image for the Spanish Radar National Composite, *Proceedings of ERAD 2006*, 318–320, 2006.
- Herr, H., Welles, E., Mullusky, M., Wu, L., Shaake, J., and Seo, D.: Probabilistic hydrologic forecasting: An ensemble approach. Second Federal Interagency Hydrologic Modelling Conference, Las Vegas, NV, 28 July–1 August, 2002 ([www.nws.noaa.gov/oh/hrl/presentations/fihm02/](http://www.nws.noaa.gov/oh/hrl/presentations/fihm02/)), 2002.
- Jurczyk, A., Osródk, K., and Szturc, J.: Research studies on improvement in real-time estimation of radar-based precipitation in Poland, *Meteorol. Atmos. Phys.*, accepted, 2008.
- Kelly, K. S. and Krzysztofowicz, R.: Precipitation uncertainty processor for probabilistic river stage forecasting, *Wat. Resour. Res.*, 36, 2643–2653, 2000.



- Krzysztofowicz, R.: Bayesian system for probabilistic river stage forecasting, *J. Hydrol.*, 268, 16–40, 2002.
- Michelson, D., Einfalt, T., Holleman, I., Gjertsen, U., Friedrich, K., Haase, G., Lindsog, M., and Jurczyk, A.: Weather radar data quality in Europe – quality control and characterization, Review, COST Action 717, Working document, Luxembourg, 2005.
- Molteni, F., Buizza, R., Palmer, T. N., and Petroliagis, T.: The ECMWF Ensemble Prediction System: Methodology and validations, *Q. J. R. Meteorol. Soc.*, 122, 73–119, 1996.
- Peura, M., Koistinen, J., and Hohti, H.: Quality information in processing weather radar data for varying user needs, *Proceedings of ERAD 2006*, 563–566, 2006.
- Pappenberger, F., Beven, K. J., Hunter, N. M., Bates, P. D., Gouweleeuw, B. T., Thielen, J., and de Roo, A. P. J.: Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS), *Hydrol. Earth Syst. Sci.*, 9, 381–393, 2005, <http://www.hydrol-earth-syst-sci.net/9/381/2005/>.
- Roulin, E. and Vannitsem S.: Skill of medium-range hydrological ensemble predictions, *J. Hydrometeor.*, 6, 729–744, 2005.
- Šálek, M., Cheze, J.-L., Handwerker, J., Delobbe, L., and Uijlenhoet, R.: Radar techniques for identifying precipitation type and estimating quantity of precipitation, COST Action 717, Working Group 1 – A review, Luxembourg, 2004.
- Sinclair, S. and Pegram, G.: Combining radar and rain gauge rainfall estimates using conditional merging, *Atmos. Sci. Lett.*, 6, 19–22, 2005.
- Szturc, J. and Dziewit, Z.: Status and perspectives on using radar data: Poland, Use of radar observations in hydrological and NWP models, COST Action 717, Final report, Luxembourg 2005, 218–221, 2005.
- Szturc, J., Ośródka, K., and Jurczyk, A.: Scheme of quality index for radar-derived estimated and nowcasted precipitation, *Proceedings of ERAD 2006*, 583–586, 2006a.
- Szturc, J., Ośródka, K., and Jurczyk, A.: Uncertainty in radar-derived estimated and nowcasted precipitation inputs to flash floods modelling, *Proceedings of ERAD 2006*, 600–602, 2006b.
- Szturc, J., Ośródka, K., and Jurczyk, A.: Parameterisation of radar precipitation quality index scheme on raingauge data, 33rd Conf. on Radar Meteorology (CD), *Proceedings. AMS*, 6–10 August 2007, Cairns, Australia, 2007a.
- Szturc, J., Ośródka, K., Jurczyk, A., and Moszkowicz, S.: Multi-sensor and multi-model case studies: from rain measurement to water level forecast and verification, RISK AWARE Project, Action 4.08 final report, Warszawa 2007 (<http://www.smr.arpa.emr.it/riskaware/get.php?file=4.08.pdf>), 2007b.
- Weipert, A. and Pierce, C.: Multi sensor supported flood detection and monitoring system for Poland. 31st Conf. on Radar Meteorology, (vol. II), *Proceedings. AMS*, 6–12 August 2003, Seattle WA, 718–721, 2003.